Variability of the Sheep Lung Microbiota

Exploring the Variability of the Sheep Lung Microbiota

Laura Glendinning[a] #, Steven Wright[a], Jolinda Pollock[a, b], Peter Tennant[a], David Collie[a], Gerry McLachlan[a]

The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, Midlothian, United Kingdom[a]; Monogastric Science Research Centre, Scotland's Rural College (SRUC), Edinburgh, Midlothian, United Kingdom [b]

#Address correspondence to Laura Glendinning, Laura.Glendinning@roslin.ed.ac.uk

Running Head: The Sheep Lung Microbiota

Abstract

Sequencing technologies have recently facilitated the characterisation of bacterial communities present in lungs during health and disease. However, there is currently a dearth of information concerning the variability of such data in health both between and within subjects. This study seeks to examine such variability using healthy adult sheep as our model system.

Protected specimen brush samples were collected from three spatially disparate segmental bronchi of six adult sheep (age 20 months) on three occasions (day 0, one month and three months). To further explore the spatial variability of the microbiota, more extensive brushings (n=16) and a throat swab were taken from a separate sheep. The V2-V3 hypervariable regions of the bacterial 16S rRNA gene were amplified and sequenced via Illumina MiSeq. DNA sequences were analysed using the MOTHUR software package. Quantitative PCR was performed to quantify total bacterial DNA.

Some sheep lungs contained dramatically different bacterial communities at different sampling sites whereas in others airway microbiota appeared similar across the lung. In our spatial variability study, clustering was observed related to the depth within the lung from which samples were taken. Lung depth refers to increasing distance from the glottis progressing in a caudal direction. We conclude that both host influence and local factors have an impact on the composition of the sheep lung microbiota.

Importance

Until recently, it was assumed that the lungs were a sterile environment which were only colonised by microbes during disease. However, recent studies using sequencing technologies

have found that there is a small population of bacteria which exists in the lung during health, referred to as the 'lung microbiota'. In this study we characterise the variability of the lung microbiota of healthy sheep. Sheep are not only economically important animals but are also often used as large animal models of human respiratory disease. We conclude that, whilst host influence does play a role in dictating the types of microbes which colonise the airways, it is clear that local factors also play an important role in this regard. Understanding the nature and influence of these factors will be key to understanding the variability in, and functional relevance of the lung microbiota.

Introduction

Within the past five years, a diverse array of bacteria has been detected in healthy lungs through the use of non-culture based methods (1, 2). These bacterial communities are commonly referred to as the lung microbiota and are thought to originate predominantly from the upper respiratory tract (3, 4). The presence of particular bacterial communities in the lung has been associated with several human diseases including cystic fibrosis (5), chronic obstructive pulmonary disease (6), bronchiectasis (7) and lung transplant rejection (8).

Whilst variation in the microbial communities present in the human lung exists at both a large and small scale, based upon the location of the bacteria within the lungs (9) and the host cell-types present (10), intra-individual variation has been found to be significantly less than inter-individual variation, indicating that each individual may play host to a specific lung microbiota (9).

57 The lung microbiota of healthy domestic sheep has previously been investigated using culture
58 based methods (11-14) but these studies have shown conflicting descriptions of the extent of
59 lung colonisation by bacteria. A study in pneumonic Bighorn sheep lungs found that for most
60 sheep studied, bacterial 16S rRNA gene amplification and sequencing was able to identify
61 additional bacterial species which were not found by culturing (15). Previous studies have also
62 examined the upper respiratory tract of healthy sheep by culture based methods (11, 12, 14, 16).
63 These studies are highly variable in the types and proportions of microbes identified.

64 Previously, our group studied the composition of the lung microbiota in sheep pre- and post-
65 infection with *Pseudomonas aeruginosa* (17). This study included the first description of the
66 lung microbiota communities of healthy domestic sheep by next generation sequencing. A
67 diverse community of microbes was identified and variability was seen to be high, both within
68 and between animals. The variability of the healthy lung microbiota at specific lung sites over
69 time has not been reported in any animal, although serial sampling of non-diseased human is
70 planned as part of the Lung HIV Microbiome Project (LHMP) (18).

71 In the present study, protected specimen brush samples were collected from three spatially
72 disparate segmental bronchi at three time-points (baseline, one month and three months) to
73 examine the composition and variability of the lung microbiota in healthy domestic sheep. In
74 addition, samples were also taken from a separate sheep from a greater number of respiratory
75 tract locations, to further explore the extent of spatial variability.

76 Such studies are fundamental to understanding the functional relevance of lung microbiota in
77 health and disease in ruminants. Indeed bacterial pneumonia is well recognised in cattle and
78 sheep and is often associated with high morbidity and mortality. Notably, regional predilection is
79 evident in that infection by *Pasteurella* occurs most frequently in the apical and cardiac lobes in

both sheep (12, 19) and cattle (20, 21). Co-infections with other respiratory pathogens are commonplace, it is already well known that infection by *Bordetella parapertussis* and *Mycoplasma ovipneumoniae* can lead to more severe disease caused by *Mannheimia (Pasteurella) haemolytica* (22-25), and there are well recognised links to stressful events such as housing or transport. As it is conceivable that changes in the lung microbiota may precipitate or associate with such events it is vital to ground future disease-related studies on a firm basis of understanding normal variation in health. Whilst the immediate focus of such studies relates to animal health it is also important to acknowledge that sheep are frequently used as models for human respiratory research (26, 27) and that there is an ongoing need to highlight any comparative contrasts and consistencies as and when they arise.

Materials and methods

Animals and airway sampling

Six twenty month old Suffolk-cross sheep were used in this study (**Table 1**) (5 females, 1 castrated male) and were housed indoors in pens for the trial duration. All animals had not undergone bronchoscopic examination during the four months preceding the study. Animal procedures were subject to the Animals (Scientific Procedures) Act 1986 and were approved by the Roslin Institute Animal Welfare and Ethics Committee.

Anaesthesia was performed as described previously (28). Sheep were sampled by protected specimen brushings (ConMed Disposable Microbiology Brush, New York, NY, USA) at 0 days (baseline), one month and three months. Sampling sites are shown in **Fig 1**. Bronchoscopy was

performed via an endotracheal tube by the same operator for all sheep at all time-points. The sample harvest dates can be found in **Table S1**. Before sampling of every sheep on any given day, 7.5 ml of phosphate buffered saline (PBS) was passed through the bronchoscope channel to act as an environmental qPCR control. Bronchoscope washings were centrifuged at 13,000 g for 15 minutes and the pellet was resuspended in 500 µl of PBS.

A throat swab and brushing samples (harvested as above) were also taken from a further sheep (female; age 36 months; 60 kg bodyweight) at a single time-point to further explore the spatial variability of the lung microbiota (sampling date 01/05/2015). Brushing sites were dorsal and ventral trachea and paired sites from either side of airway bifurcations progressing along the anterior to caudal lung axis (**Fig 2**).

DNA extraction, amplification and sequencing

DNA extraction was performed using the MOBIO PowerSoil® DNA Isolation Kit (Carlsbad, CA, USA). Brushes were transferred into PowerSoil Bead Tubes with PowerSoil Solution C1 and PowerSoil Bead Solution. Bead Tubes were heated at 65°C for 10 minutes then placed in a FastPrep FP120 Cell Disrupter (Qbiogene Inc., Cedex, France) for 45 seconds at 5.0 m/sec. From this point onwards the manufacturer's instructions were followed, except for the final elution step. Purified DNA was eluted into 50 µl of PowerSoil Solution C6 rather than 100 µl to increase the DNA concentration.

All PCR steps used Q5® High-Fidelity 2X Master Mix (New England Biolabs, Beverly, MA, USA). A nested PCR reaction was performed with Illumina adaptor sequences and barcodes (**Table S2**) included only on the primers for the second round in an attempt to reduce bias

caused by barcoded primers when amplifying low biomass samples (29). The conditions for the first round of PCR, amplifying the V1-V4 16S hypervariable regions (primers: 28F (5'–GAGTTTGATCNTGGCTCAG–3') and 805R (5'–GACTACCAGGGTATCTAATC–3')), were: 94°C for 2 minutes followed by 20 cycles of 94°C for 1 minute, 55°C for 45 seconds and 72°C for 1.5 minutes followed by 72°C for 20 minutes. The conditions for the second round of PCR, amplifying the V2-V3 16S hypervariable regions (primers: 104F (5'–GGCGVACGGGTGAGTAA–3') and 519R (5'–GTNTTACNGCGGCKGCTG–3')), were: 98°C for 30 seconds followed by 20 cycles of 98°C for 10 seconds, 67°C for 30 seconds and 72°C for 10 seconds followed by 72°C for 2 minutes. Amplicons from both rounds of PCR were purified using the AMPure XP PCR Purification System (Beckman Coulter, Brea, CA, USA). Amplicons were sequenced using an Illumina MiSeq or Hiseq (Illumina, San Diego, CA) run producing paired end 250-nucleotide reads (30). Those samples sequenced by two Miseq runs are listed in Dataset S1 and those sequenced by Hiseq are listed in Dataset S2. Where samples from the Miseq runs were found to have low read numbers, they were sequenced again on a separate Miseq run (Samples: 2D618 RA 3 months and 2D619 RA 3 months). We previously confirmed cross-run stability by comparing separate runs made on the same samples (**Fig S1**).

Extraction kit controls were produced by carrying out a reagent-only extraction using the MOBIO PowerSoil® DNA Isolation Kit. PCR reagent controls were constructed by adding 20 µl of nuclease free water to the PCR reaction mixture. The Human Microbiome Project Mock Community HM-782D (100,000 copies per organism per µl, BEI Resources, ATCC, Manassas, VA, USA) , extraction kit controls, PCR reagent-only controls and positive controls (DNA extracted from *Pseudomonas aeruginosa* Strain PA0579) were amplified and sequenced by the same methods as were used for samples.

A separate mock community sample was sequenced using an Illumina HiSeq. For this sample, the solution produced from the first round of PCR was diluted 1:100 in nuclease free water before being used in the second round of PCR. This was carried out to ascertain the effect of placing different concentrations of DNA into the second PCR round on PCR bias.

The unassembled reads, with primers removed, are publicly available through the NCBI Sequence Read Archive (SRA) under the Bioproject ascension: PRJNA298882.

Bioinformatic and statistical analysis

Primers were removed using Cutadapt (31). Sequences which contained more than one base error per 10 primer bases were removed from further analysis. The following steps were carried out in MOTHUR (32) and were based upon a protocol developed for MiSeq by the MOTHUR creators (30). Forward and reverse reads were aligned to form one continuous DNA sequence; any sequences which failed to align were discarded. Sequences which contained ambiguous bases; were less than 369 base pairs in length or contained homopolymers of greater than 9 base pairs were also discarded. Chimeras were identified and removed using UCHIME (33). Sequences were aligned to the SILVA reference alignment (34) and were classified using MOTHUR's Bayesian classifier against the Greengenes database (35), which was trimmed to the V2-V3 hypervariable region of the 16S rRNA gene to improve classification depth (36). Sequences identified as not originating from bacteria were removed from further analysis. Operational taxonomic units (OTUs) were clustered into phylotypes using a database-dependent approach then sub-sampled.

168 Distance matrices were created using Yue and Clayton theta values (37). Analysis of molecular
169 variance (AMOVA) (38) was used to determine significant differences between the bacterial
170 compositions of groups. Principal co-ordinate analysis (PCOA) graphs were constructed to
171 visualise similarities between samples. The Inverse Simpson's index was used to quantify
172 diversity. Where data was non-parametric the Friedman test was used to identify significant
173 differences in diversity, using Minitab® 16 for Windows (Minitab, Coventry, UK). All other
174 statistical tests were carried out within MOTHUR. Metastats (39) was used to identify OTUs
175 which were different between groups. Good's coverage (40) was used to estimate sample
176 coverage and the Chao 1 index was used to calculate richness. Indicator OTUs (OTUs which are
177 indicative of a particular group of samples) were identified using the indicator metric within
178 MOTHUR (41). Repeated measures ANOVAs were carried out using the Vegan package in R
179 (42-44).

180

181 **qPCR**

182 qPCR reactions were performed using the LightCycler® 480 SYBR Green I Master Mix (Roche
183 Applied Science, Indianapolis, IN, USA), 1 µl of extracted DNA solution and the 16S rDNA
184 qPCR primers UniF340 (5'–ACTCCTACGGGAGGCAGCAGT–3') and UniR514 (5'–
185 ATTACCGCGGCTGCTGGC–3') at a final concentration of 0.4 µM.

186 The qPCR run consisted of a pre-incubation step of 50°C (ramp rate: 4.80°C/s for 2 minutes)
187 then 95°C (ramp rate: 4.80°C/s for 10 seconds) and an amplification step consisting of 45 cycles
188 of 95°C (ramp rate: 4.80°C/s for 30 seconds) then 63°C (ramp rate: 2.50°C/s for 30 seconds).
189 This was followed by a melting cycle consisting of 95°C (ramp rate: 4.80°C/s for 5 seconds) then

65°C (ramp rate: 4.80°C/s for 1 minute) followed by 97°C (ramp rate: 0.11°C/s, acquisition mode, continuous).

Negative controls consisted of both water and extraction kit reagent controls. For water controls, 1 µl of nuclease free water was added to the qPCR reaction mixture. For extraction kit controls, DNA extractions were carried out using the MOBIO PowerSoil® DNA Isolation Kit (Carlsbad, CA, USA) following the same protocol as was used to extract DNA from samples, except no sample was added meaning that any bacterial DNA in the final elution must have been derived from the extraction kit reagents. 1 µl of this elution was added to the qPCR reaction mixture.

In order for us to compare the quantity of bacterial DNA found in bronchoscope wash and brushing samples it was necessary to use a unit of measurement which could be applied to both sample types. Bacterial DNA concentrations are therefore reported as the 16S copy numbers present per µl of eluent produced from samples by the MOBIO PowerSoil® DNA Isolation Kit. Statistical analysis was carried out in Minitab® 16 for Windows. Where data was non-parametric the Mann-Whitney U test was used to statistically compare groups.

Results

Quality control and adequacy of sequencing

After constructing DNA sequences from the forward and reverse reads generated by sequencing, various quality control steps were performed to decrease the number of artefacts and poor quality sequences used in subsequent analyses.

For the Miseq runs, these steps resulted in a 15% loss of sequences (sequencing error rate = 0.39%). On average, samples contained 205625 ± 27232 (mean ± standard error of the mean (SEM)) sequences and a total of 925 bacterial OTUs were identified (**Dataset S1**). Sequences were assigned to OTUs based on their taxonomic classifications. Each OTU does not necessarily represent an individual bacterial species but instead represents the lowest taxonomic level to which its bacterial sequences could be assigned. For example, 77.4% of reads could be identified to genus while 31.1% could be assigned to species. If two species from the same genus could only be assigned to genus level then they would both be binned into the same OTU.

For the Hiseq run, samples contained on average 233505 ± 69735 (mean ± SEM) and the sequencing error rate was 0.39%. 633 OTUs were identified (**Dataset S2**) and the total reduction in sequence numbers due to quality control was 5%.

Good's coverage estimate values exceeded 97% for all samples. This indicates that at least 97% of the bacteria present in our original samples are likely to have been identified, demonstrating that the depth of sequencing was adequate.

Of the twenty bacteria contained in the mock community, all could be taxonomically identified down to genus level except *Bacillus cereus*, *Escherichia coli* and *Listeria monocytogenes* which could only be identified at family level. This indicates that the primers were able to amplify a wide diversity of bacteria. Whilst the proportions of bacterial DNA were different to the proportions anticipated if no PCR bias was present (**Table 2**) this was less apparent in the sample which had been diluted 1:100 after the first round of PCR. In the undiluted mock the proportions of bacterial orders differed from the expected proportions by an average of 9.48% (SEM: 2.24%, range: 0.99% - 19.48%) whereas the orders in the diluted mock differed on average by 4.33% (SEM: 1.12%, range: 0.29% - 12.71%). This diluted mock community may be more comparable

234 to the kind of biases we would find in our samples as the undiluted mock community contained a
235 far higher concentration of template DNA (2000000 16S copies per µl) than our samples did on
236 average (13133 16S copies per µl).

237 The assumption was made that PCR bias would reasonably be expected to apply equally across
238 all samples and therefore any statistical tests between samples should still be valid. The two
239 bacterial species most overrepresented in the undiluted mock community (*Deinococcus*
240 *radiodurans* and *Helicobacter pylori*) are not commonly associated with the respiratory tract and
241 bacteria from these genera were very rare within our dataset.

242

243 **Longitudinal study in six sheep over three months**

244 To examine the spatial, longitudinal and inter-individual variation of the sheep lung microbiota,
245 lung brushings were taken from three spatially disparate lung locations (RA, RCD and LCD) in
246 six sheep at three time-points (baseline, one month and three months). Estimates of total bacterial
247 yield from qPCR analysis indicated that sheep lung brushings contained an average of 13133 ±
248 894 (mean ± SEM) 16S copy numbers/µl (range: 1032-37627 16S copy numbers/µl).
249 Bronchoscope wash controls contained significantly lower bacterial 16S rDNA concentrations
250 than lung brushings (Mann-Whitney U test: $p < 0.0001$), containing an average of 1471 ± 279
251 (mean ± SEM) 16S copy numbers/µl (range: 397-4792 16S copy numbers/µl) (**Fig 3**). The qPCR
252 negative water controls were found to contain 190, 479 and 739 16S copy numbers/µl and the
253 extraction kit controls were found to contain 347 and 511 16S copy numbers/µl.

254 After sequencing and sub-sampling, bacterial communities isolated from extraction kit and 16S
255 PCR negative controls were found to cluster separately from those found in sheep lung brushings

(AMOVA: $p < 0.001$). Extraction kit controls were included from two different lots. The most abundant OTUs found in the first extraction kit control were *Corynebacterium* (36%), Enterobacteriaceae (13%), *Mycobacterium llatzerens* (7%) and *Staphylococcus haemolyticus* (5%). The most predominant OTUs in the second extraction kit control were *Aerococcus* (13%), Dermabacteraceae (11%), *Micrococcus* (10%), *Enhydrobacter* (9%) and *Leuconostoc* (7.2%). The predominant bacterial order present in both extraction kit controls was Actinomycetales (50.1% and 40.5% respectively).

The bacteria isolated from lung brushings predominantly belonged to the orders Bacillales (26%), Actinomycetales (21%), Clostridiales (11%) and Lactobacillales (9%) while common genera included *Staphylococcus* (16%), *Corynebacterium* (9%), *Jeotgalicoccus* (5%) and *Streptococcus* (5%).

The underlying changes in bacterial OTUs between sampling points were examined. The bacterial communities found in lung brushings clustered significantly by time-point (AMOVA: $P < 0.001$) (**Fig 4**). The OTUs causing this clustering were identified by applying Metastats (**Tables S3 and S4**). The largest difference observed between the first and second time-points was an 11% increase in the abundance of an OTU identified as *Corynebacterium*. This is also the most abundant OTU in one of our extraction kit controls. OTU 12: *Mycobacterium llatzerense* was also significantly more abundant at the one month time-point and was the third most abundant OTU in the same extraction kit control. It is therefore likely that our time-points were affected to different degrees by reagent contamination and therefore the analysis of segments over time is not possible. However, all samples taken in the same sheep at the same time-point were processed using the same extraction kit; therefore, an analysis of spatial variability can be performed.

Visual perceptions of community structure indicated that in some sheep, samples taken from separate lung sites differed appreciably, whereas in other sheep there appeared relative concordance between such samples (example shown in **Fig 5**). A full visual summary of the results can be found in **Fig S2**. There were no significant differences between the diversity of communities located at different lung sites (Inverse Simpson's index: Friedman test: $P > 0.5$).

Sheep clustered separately by the composition of their lung bacterial communities at the baseline time-point (AMOVA: $P = 0.001$), and at the three month time-point (AMOVA: $P = 0.045$), indicating that samples taken from within the same sheep were more similar to one another than to samples taken from other sheep. At the one month time-point, sheep did not cluster in this manner (AMOVA: $P = 0.394$), though this is likely due to the presence of contamination causing a homogenisation of our one month samples. Pairwise comparisons of samples showed no significant results. The similarity of samples to one another can be visualised using PCOA graphs (**Fig 6**).

Spatial variability of the lung microbiota in an individual sheep

The observed variability between spatially disparate lung sites in some sheep prompted enquiry as to the consistency of bacterial communities sampled from sites in close spatial apposition.

Further samples were derived by systematically sampling multiple sites of the lungs of an individual animal at one time-point. Whilst the three month experiment did not include a control for every lot of extraction kit used, emerging literature and opinion within the field has since

indicated the value of using the same extraction kit for all samples. This strategy was therefore adopted for these latter samples which were all processed at the same time.

The extraction kit control was mainly composed of one OTU (OTU 18: 79%) which was also present in our brushing samples (mean ± SEM: 51.1 ± 3.3). We felt confident in removing this OTU from all of our samples prior to analysis as it could be identified to species level (*Methylobacterium komagatae*) and was considered highly unlikely to be found within the sheep lung. No further OTUs were removed before analysis.

Lung brushings contained on average 2116 16S copy numbers per ul (SEM = 365) while the throat swab and extraction kit control contained 42480 and 43 16S copy numbers per ul respectively. The richness and diversity of the lung samples (Chao = 103.77 ± 7.32, Inverse Simpson's index = 14.24 ± 2.14) was found to be far lower than in the throat swab (Chao = 257.038, Inverse Simpson's index = 9.19). Sample A1, taken from the ventral aspect of the trachea just caudal to the bifurcation with the right apical lobe segmental bronchus, had the second highest richness (Chao = 155.024) and diversity (Inverse Simpson's index = 8.713). However, sample A2, which was taken at the same level as sample A1 but from the dorsal aspect of the trachea, had a much lower richness (Chao = 76.038) and diversity (Inverse Simpson's index = 4.925).

The composition of the communities taken from the respiratory tract showed some variation, even between paired samples located very closely to one another (**Fig 7**). Sub-tracheal samples paired to their most proximate neighbour did not cluster together significantly when OTUs were defined at the lowest taxonomic depth (AMOVA: P = 0.30). However, paired samples did cluster significantly by the bacterial orders which they contained (AMOVA: P = 0.046). Sub-tracheal samples also clustered significantly (by order) based upon the depth in the lung from which

samples were taken (AMOVA: P = 0.033) (**Fig 8**) (lung depth in this context refers to increasing distance from the glottis progressing in a caudal direction). An indicator OTU for the group which included the samples A4, A5, A14 and A15 was found to be OTU 4: Pseudomonadales (P= 0.042). The most abundant bacterial orders identified from brushings were Clostridiales (25.8%), Pseudomonadales (18.3%) and Actinomycetales (16.0%) while the throat swab was dominated by Pasteurellales (36.5%) and Pseudomonadales (15.1%). The extraction kit control was predominantly composed of Actinomycetales (31.1%) and Pseudomonadales (31.0%).

As the Pasteurellales order contains several species which are known to act as sheep lung pathogens and which display regional patterns of infection we felt it would be interesting to investigate where OTUs belonging to this order were found within the respiratory tract (**Table 3**). By far the largest proportion of these OTUs was found in the throat swab and one of the tracheal brushings (A1).

Discussion

In order to better understand the variability present in the sheep lung microbiota, we compared the lung bacterial communities of six sheep at three different lung sites over a duration of three months. To further explore the extent of spatial variability, we also took 17 samples from the respiratory tract of one sheep.

Previously, the bacteria in healthy domestic sheep lungs had been investigated by culture-based methods which seemed to indicate that bacterial colonisation of the sheep lung was rare or did not occur in all sheep (11, 12, 14). In contrast, using non-culture based methods we have found

Applied and Environmental Microbiology

345 that all of the sampled sites in our seven sheep harboured diverse communities of bacteria,
346 although in far smaller numbers than is generally found in other niches such as the gut or upper
347 respiratory tract.

348 Bacteria belonging to genera previously isolated from goat and sheep lungs (11, 12) were found
349 in our samples. These included *Corynebacterium*, *Bacillus*, *Enterococcus*, *Klebsiella*,
350 *Mannheimia*, *Micrococcus*, *Moraxella*, *Pasteurella*, *Pseudomonas*, *Staphylococcus* and
351 *Streptococcus*. Of the most common genera observed within our animals *Staphylococcus,*
352 *Streptococcus* and *Corynebacterium* are commonly isolated from the upper respiratory tract and
353 skin of many animals whereas *Jeotgalicoccus* is a less well known genus (45) which has not
354 been found to make up a substantial part of the lung microbiota communities in any previous
355 studies. However, it has been isolated from the small intestinal mucosa of calves (46), the canine
356 oral cavity (47), aerosols in a poultry house (48, 49), cattle teats (50), lamb meat (51), the rumen
357 of cattle (52) and from aerosol samples near a dairy (53).

358 The most common bacterial orders found in the sheep lung during the three month study were
359 Bacillales, Actinomycetales and Clostridiales. This agrees with the findings of a previous study
360 carried out by our group which examined the sheep lung microbiota before and after infection
361 with *P.aeruginosa* (17). Pseudomonadales (mainly *Pseudomonas*) was also commonly found in
362 the lungs during our single sheep study while the throat swab from this study was dominated by
363 Pasteurellales and Pseudomonadales.

364 Co-infection with *Bordetella parapertussis* or *Mycoplasma ovipneumoniae* has been shown to
365 lead to more severe disease caused by *Mannheimia (Pasteurella) haemolytica* (22-25).
366 Mycoplasmas were very rare within our dataset with only one sheep segment containing reads
367 from this genus at one time-point. We did not identify any OTUs as *Bordetella*; however, we did

find an OTU designated as Alcaligenaceae (the family to which *Bordetella* belongs) though these were uncommon and occurred in low abundance. We identified several OTUs which were classified as members of the Pasteurellaceae family including *Mannheimia* and *Bibersteinia* and less commonly: *Aggregatibacter segnis*, *Haemophilus parainfluenzae*, *Bibersteinia trehalosi* and *Actinobacillus parahaemolyticus*. All of these microbes have previously been isolated from the lungs or upper respiratory tract (54-58). Despite the fact that disease by members of this family is often located in the apical and cardiac lobes (12, 19), we observed members of this family to be present across the lung.

The composition of the lung microbiota found in our sheep shows some differences to that previously identified in humans where Bacteroidales are found in higher numbers and there are generally less members of the Actinomycetales and Clostridiales (2, 9, 59). Segal *et al*. identified various bacterial taxa that were commonly found in high relative abundance in human lungs (1). These included taxa which were found in all of our sheep samples in high relative abundance (*Streptococcus*, *Staphylococcus*, *Corynebacterium*); taxa which were found in the majority of our samples but in lower abundances (*Propionibacterium*, *Pseudomonas*) and taxa which were only found sporadically in our samples and were usually in low abundance (*Stenotrophomonas*, *Prevotella*, *Veillonella*, *Fusobacterium*, *Porphyromonas*).

Such differences may at least in part reflect the different surroundings in which sheep live as well as behavioural or physiological features such as rumination. A study using buccal swabs to identify bacteria originating from the rumen suggested that as the time between regurgitation and sampling increases the orally associated bacterial populations in the buccal cavity will increase and the rumen associated bacteria will decrease, potentially contributing to inter-animal variation

(58). In future studies it may be useful to take rumen and upper respiratory tract samples alongside lung samples to explore if the variability between these sites and the lung is related.

Regardless of the highlighted differences between sheep and the human lung microbiota, there is a pressing need to understand the mechanisms that underlay the spatial and temporal variability of microbiota in the mammalian lung. These fundamental studies are difficult to facilitate in healthy human subjects as a consequence of the invasive nature of the repeated sampling protocol, as well as controlling for the influence of environmental and/or lifestyle factors. Large animal models can however play an important role in filling this need. Indeed, the physiological and immunological similarities between sheep and human lungs (60, 61) have contributed to the widespread use of sheep as translational models for human lung research (26, 27) including asthma (62-65), the delivery of drugs via the upper respiratory tract (66-68), emphysema (69-71), pulmonary hypertension (72-74), physical lung injury (75-78), lung infection (28, 79-81), respiratory distress syndromes (82-85), asbestosis (86-88) and lung cancer (89, 90).

In our study we examined the variability of the lung microbiota in sheep. Bacterial populations were often different between lung segments and between individuals, which confirms our previous observations (17). There was more similarity between samples from the same sheep at the baseline and three month time-points than between samples taken from different sheep but this was not found to be the case at the one month time-point. Lung sample clustering by individuals has previously been identified in humans (9) and sheep (17).

Clearly, large differences can exist in the microbiota sampled from different lung segments at the same time-point. This spatial variability of lung microbial populations can be observed in *P.aeruginosa* infections in cystic fibrosis patient lungs (91). The mechanisms underlying such observations have yet to be elucidated; however, possible candidate influences may include

Applied and Environmental Microbiology

413 regional variability of physiological parameters such as gas concentrations, osmolality,
414 temperature, pH and blood flow (92-96) which may lead to the creation of 'microhabitats'
415 providing a selective advantage to certain bacteria (97). It has previously been demonstrated that
416 differences in pH can lead to changes in the colonic microbiota (98) and that temperature
417 combined with humidity can lead to changes in the composition of the skin microbiota (99).

418 A longitudinal analysis of the lung microbiota at specific lung sites in healthy individuals has not
419 previously been reported. Our goal was to define the variability of the lung microbiota over time
420 and to detect whether there was a sheep lung microbiota 'signature' which remains stable.
421 Unfortunately, at the time of carrying out this study the extent of the variability of bacterial DNA
422 found within different lots of extraction kits was not yet known (100). While we therefore did
423 include some extraction kit controls for our longitudinal study we did not include controls for all
424 lots which were used. Samples from different time-points were also processed at different times.
425 Due to our small sample sizes and the fact that samples clustered significantly by time-point, we
426 do not feel that accurate conclusions can be drawn about the temporal stability of the microbiota
427 from our data. However, all samples taken from the same time-point in the same sheep were
428 processed at the same time. Therefore, we can be confident that the spatial variability we observe
429 within animals is not due to our methodology.

430 In some individuals, samples taken from different lung segments were found to be highly
431 different from one another whereas in others the lung microbiota appeared to be quite stable
432 across the lung. Another finding was the disappearance of the significantly separate clustering of
433 sheep samples at the one month time-point. This was correlated with an increase in the
434 proportions of several OTUs found in sheep lungs, the most noticeable increase arising from an
435 OTU classified as *Corynebacterium* which was also the most abundant OTU in one of our

extraction kit controls. It is likely that the disappearance of significant clustering by individual at the one month time-point is due to the increased presence of contamination in our samples.

OTUs that were identified in both samples and negative PCR and extraction kit controls were not removed from the analysis for the three month sheep study. The reason for this decision was that a number of bacteria commonly associated with the upper and lower respiratory tract were present in these controls, including the genera *Streptococcus* and *Pseudomonas*, and it was judged that their removal would merely introduce another source of bias.

Equally, any specific *a priori* manipulation based around assumptions gleaned from the human literature regarding microbiota in the upper and lower respiratory tract, are potentially ill advised. Indeed, it has been demonstrated that the microbes found in the lungs of animals often match those found in their bedding and hay (101). It is therefore not possible to dismiss environmental microorganisms as being only due to the contamination of samples.

In our spatial variation study, one OTU was removed before analysis as we felt confident that its presence was due to contamination of our extraction kit. Clustering of lung brushings by the lung depth from which they were taken was observed when OTUs were defined by bacterial order. Samples paired with their proximate neighbour were also found to cluster significantly separately from brushings taken elsewhere in the lung but this may just be due to the fact that these samples were taken from the same lung depth. Certainly further research to explore the relationship between lung depth and community composition appears warranted.

After sequencing a mock community of bacteria which contained equimolar concentrations of each bacterial species we did find some bias present, with some bacterial species being over or underrepresented. These biases, which may be caused by various factors including primer

mismatching, PCR cycle number and the bioinformatic pipeline used, are quite common in 16S sequencing (102-105). We also sequenced a 1:100 dilution of the same mock community and found that the apparent biases were far less. As the concentration of bacterial DNA in our samples was far lower than that of the undiluted mock community, we feel that the 1:100 dilution is likely to better represent the biases which may be present in our samples as it is closer to their bacterial DNA concentrations. We believe that this vindicates our choice of DNA amplification strategy, including the use of nested PCR.

It may not be possible to claim that the bacterial abundances identified via 16S sequencing quantitatively represent the relative abundances of bacteria in the sample. Indeed, this is made even more difficult as different bacterial taxa contain different copy numbers of the 16S gene (106). However, it seems logical to assume that if the same methodology is used for all samples within a study then the biases present will be the same for all samples and therefore comparisons between groups or claims about the types of microbes present in samples would still be valid.

In conclusion, we observed variability in the sheep lung microbiota both between and within individuals. In some animals different lung segments contained highly different bacterial communities whereas other animals showed similar communities at all lung sites. While spatial variation was observed to occur over both large and small distances across the lung, samples taken at the same lung depth clustered together separately from those taken at different lung depths. Further studies are needed to explore the stability of the healthy lung microbiota over time.

Funding Information

Funding for this study was supplied by the BBSRC (Biotechnology and Biological Sciences Research Council) under the EASTBIO doctoral training program (Grant number: BB/J01446X/1). The funders had no input on study design; data collection and interpretation or the decision to submit this study for publication.

Acknowledgments

The authors would like to thank the administrative staff of the BBSRC EASTBIO DTP program for their encouragement and support. We would like to thank the staff at Dryden farm for providing the everyday care of our animals. We would also like to thank Dr. Helen Brown for her advice on statistical methods and sample sizes and Tina Baker for providing the qPCR protocol. The Human Microbiome Project Mock Community HM-782D was obtained through BEI Resources, NIAID, NIH as part of the Human Microbiome Project: Genomic DNA from Microbial Mock Community B (Even, Low Concentration), v5.1L, for 16S rRNA Gene Sequencing, HM-782D.

References

1. **Segal L, Alekseyenko A, Clemente J, Kulkarni R, Wu B, Chen H, Berger K, Goldring R, Rom W, Blaser M, Weiden M.** 2013. Enrichment of lung microbiome with supraglottic taxa is associated with increased pulmonary inflammation. Microbiome **1:**19.

2. **Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, Bushman FD, Collman RG.** 2011. Topographical continuity of bacterial populations in the healthy human respiratory tract. Am J Respir Crit Care Med **184:**957-963.

3. **Morris A, Beck JM, Schloss PD, Campbell TB, Crothers K, Curtis JL, Flores SC, Fontenot AP, Ghedin E, Huang L, Jablonski K, Kleerup E, Lynch SV, Sodergren E, Twigg H, Young VB, Bassis CM, Venkataraman A, Schmidt TM, Weinstock GM.** 2013. Comparison of the respiratory microbiome in healthy nonsmokers and smokers. Am J Respir Crit Care Med **187:**1067-1075.

4. **Charlson E, Bittinger K, Chen J, Diamond J, Li H, Collman R, Bushman F.** 2012. Assessing bacterial populations in the lung by replicate analysis of samples from the upper and lower respiratory tracts. PLoS ONE **7:**e42786. DOI: 10.1371/journal.pone.0042786

5. **Delhaes L, Monchy S, Fréalle E, Hubans C, Salleron J, Leroy S, Prevotat A, Wallet F, Wallaert B, Dei-Cas E, Sime-Ngando T, Chabé M, Viscogliosi E.** 2012. The airway microbiota in cystic fibrosis: A complex fungal and bacterial community—implications for therapeutic management. PLoS ONE **7:**e36313. DOI: 10.1371/journal.pone.0036313

6. **Erb-Downward JR, Thompson DL, Han MK, Freeman CM, McCloskey L, Schmidt LA, Young VB, Toews GB, Curtis JL, Sundaram B, Martinez FJ, Huffnagle GB.** 2011. Analysis of the lung microbiome in the "healthy" smoker and in COPD. PLoS ONE **6:**e16384. DOI: 10.1371/journal.pone.0016384

7. **Boyton RJ, Reynolds CJ, Quigley KJ, Altmann DM.** 2013. Immune mechanisms and the impact of the disrupted lung microbiome in chronic bacterial lung infection and bronchiectasis. Clin Exp Immunol **171:**117-123.

8. **Borewicz K, Pragman AA, Kim HB, Hertz M, Wendt C, Isaacson RE.** 2013. Longitudinal analysis of the lung microbiome in lung transplantation. FEMS Microbiol Lett **339:**57-65.

9. **Dickson RP, Erb-Downward JR, Freeman CM, McCloskey L, Beck JM, Huffnagle GB, Curtis JL.** 2015. Spatial variation in the healthy human lung microbiome and the adapted island model of lung biogeography. Ann Am Thorac Soc **12:**821-830.

10. **Dickson R, Erb-Downward J, Prescott H, Martinez F, Curtis J, Lama V, Huffnagle G.** 2014. Cell-associated bacteria in the human lung microbiome. Microbiome **2:**28.

11. **Yimer N, Asseged B.** 2007. Aerobic bacterial flora of the respiratory tract of healthy sheep slaughtered in Dessie municipal abattoir, northeastern Ethiopia. Rev Med Vet Toulouse **158:**473-478.

12. **Alley MR.** 1975. The bacterial flora of the respiratory tract of normal and pneumonic sheep. New Zeal Vet J **23:**113-118.

13. **Alley MR, Quinlan JR, Clarke JK.** 1975. The prevalence of *Mycoplasma ovipneumoniae* and *Mycoplasma arginin*i in the respiratory tract of sheep. New Zeal Vet J **23:**137-141.

14. **Barbour EK, Nabbut NH, Hamadeh SK, Al-Nakhli HM.** 1997. Bacterial identity and characteristics in healthy and unhealthy respiratory tracts of sheep and calves. Vet Res Commun **21:**421-430.

15. **Besser TE, Cassirer EF, Potter KA, VanderSchalie J, Fischer A, Knowles DP, Herndon DR, Rurangirwa FR, Weiser GC, Srikumaran S.** 2008. Association of *Mycoplasma ovipneumoniae* infection with population-limiting respiratory disease in free-ranging Rocky Mountain Bighorn Sheep (*Ovis canadensis canadensis*). J Clin Microbiol **46:**423-430.

16. **Queen C, Ward ACS, Hunter DL.** 1994. Bacteria isolated from nasal and tonsillar samples of clinically healthy Rocky Mountain bighorn and domestic sheep. J Wildl Dis **30:**1-7.

17. **Collie D, Glendinning L, Govan J, Wright S, Thornton E, Tennant P, Doherty C, McLachlan G.** 2015. Lung microbiota changes associated with chronic *Pseudomonas aeruginosa* lung infection and the impact of intravenous colistimethate. PloS ONE **10:**e0142097. DOI: 10.1371/journal.pone.0142097

18. **Dickson RP, Erb-Downward JR, Huffnagle GB.** 2013. The role of the bacterial microbiome in lung disease. Expert Rev Respir Med **7:**245-257.

19. **Davies DH, Dungworth DL, Humphreys S, Johnson AJ.** 1977. Concurrent infection of lambs with parainfluenza virus type 3 and *Pasteurella-haemolytica*. New Zeal Vet J **25:**263-265. 10.1080/00480169.1977.34425

20. **Catt DM, Chengappa MM, Kadel WL, Herren CE.** 1985. Preliminary studies with a live streptomycin dependant *Pasteurella-multocida* and *Pasteurella-haemolytica* vaccine for the prevention of bovine pneumonic pasteurellosis. Can J Vet Res **49:**366-371.

21. **Stockdale PHG, Langford EV, Darcel CLQ.** 1979. Experimental bovine pneumonic paseturellosis.1. Prevention of the disease. Can J Comp Med **43:**262-271.

22. **Brogden KA, Lehmkuhl HD, Cutlip RC.** 1998. *Pasteurella haemolytica* complicated respiratory infections in sheep and goats. Vet Res **29:**233-254.

23. **Dassanayake RP, Call DR, Sawant AA, Casavant NC, Weiser GC, Knowles DP, Srikumaran S.** 2010. *Bibersteinia trehalosi* inhibits the growth of *Mannheimia haemolytica* by a proximity-dependent mechanism. Appl Environ Microbiol **76:**1008-1013.

24. **Dassanayake RP, Shanthalingam S, Herndon CN, Subramaniam R, Lawrence PK, Bavananthasivam J, Cassirer EF, Haldorson GJ, Foreyt WJ, Rurangirwa FR, Knowles DP, Besser TE, Srikumaran S.** 2010. *Mycoplasma ovipneumoniae* can predispose bighorn sheep to fatal *Mannheimia haemolytica* pneumonia. Vet Microbiol **145:**354-359.

25. **Porter JF, Connor NK, Krueger N, Hodgson JC, Donachie W.** 1995. Predisposition of specific pathogen-free lambs to *Pasteurella haemolytica* pneumonia by *Bordetella parapertussis* infection. J Comp Pathol **112:**381-389.

26. **Scheerlinck JPY, Snibson KJ, Bowles VM, Sutton P.** 2008. Biomedical applications of sheep models: From asthma to vaccines. Trends Biotechnol **26:**259-266.

27. **Meeusen EN, Snibson KJ, Hirst SJ, Bischof RJ.** 2009. Sheep as a model species for the study and treatment of human asthma and other respiratory diseases. Drug Discov Today Dis Models **6:**101-106.

28. **Collie D, Govan J, Wright S, Thornton E, Tennant P, Smith S, Doherty C, McLachlan G.** 2013. A lung segmental model of chronic *Pseudomonas* infection in sheep. PLoS ONE **8:**e67677. DOI: 10.1371/journal.pone.0067677

29. **Berry D, Ben Mahfoudh K, Wagner M, Loy A.** 2011. Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. Appl Environ Microbiol **77:**7846-7849.

30. **Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD.** 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq illumina sequencing platform. Appl Environ Microbiol **79:**5112-5120.

31. **Martin M.** 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. Embnetjournal **17:**10-12.

32. **Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF.** 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol **75:**7537-7541.

33. **Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R.** 2011. UCHIME improves sensitivity and speed of chimera detection. BMC Bioinform **27:**2194-2200.

34. **Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO.** 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res **41:**D590-D596. DOI: 10.1093/nar/gks1219

35. **DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL.** 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol **72:**5069-5072.

36. **Werner JJ, Koren O, Hugenholtz P, DeSantis TZ, Walters WA, Caporaso JG, Angenent LT, Knight R, Ley RE.** 2012. Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys. ISME J **6:**94-103.

37. **Yue JC, Clayton MK.** 2005. A similarity measure based on species proportions. Commun Stat-Theory Methods **34:**2123-2131.

38. **Excoffier L, Smouse PE, Quattro JM.** 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics **131:**479-491.

39. **Paulson J, Pop M, Bravo H.** 2011. Metastats: an improved statistical method for analysis of metagenomic data. Genome Biol **12 (Supplement 1):**27.

40. **Esty WW.** 1986. The efficiency of Good's nonparametric coverage estimator. Ann Stat **14:**1257-1260.

41. **Dufrêne M, Legendre P.** 1997. Species assemblages and indicator species:The need for a flexible assymetrical approach. Ecol Monogr **67:**345-366.

42. **Efron B, Tibshirani RJ.** 1993. An introduction to the bootstrap, monographs on statistics and applied probability, Vol. 57. Taylor and Francis, Abigndon, UK.

43. **Dixon P.** 2003. VEGAN, a package of R functions for community ecology. J Veg Sci **14:**927-930.

44. **Anderson M, Gorley RN, Clarke RK.** 2008. Permanova+ for primer: Guide to software and statistical methods. PRIMER-E, Plymouth, UK.

45. **Schwaiger K, Hölzel C, Mayer M, Bauer J.** 2010. Notes on the almost unknown genus *Jeotgalicoccus*. Lett Appl Microbiol **50:**441-444.

46. **Malmuthuge N, Griebel PJ, Guan LL.** 2014. Taxonomic identification of commensal bacteria associated with the mucosa and digesta throughout the gastrointestinal tracts of preweaned calves. Appl Environ Microbiol **80:**2021-2028.

624 47. **Dewhirst FE, Klein EA, Thompson EC, Blanton JM, Chen T, Milella L, Buckley CMF, Davis IJ,**
625 **Bennett M-L, Marshall-Jones ZV.** 2012. The canine oral microbiome. PLoS ONE **7:**e36067.
626 DOI: 10.1371/journal.pone.0036067
627 48. **Fallschissel K, Klug K, Kämpfer P, Jäckel U.** 2010. Detection of airborne bacteria in a German
628 turkey house by cultivation-based and molecular methods. Ann Occup Hyg **54:**934-943.
629 49. **Martin E, Fallschissel K, Kämpfer P, Jäckel U.** 2010. Detection of *Jeotgalicoccus* spp. in
630 poultry house air. Syst Appl Microbiol **33:**188-192.
631 50. **Braem G, De Vliegher S, Verbist B, Piessens V, Van Coillie E, De Vuyst L, Leroy F.** 2013.
632 Unraveling the microbiota of teat apices of clinically healthy lactating dairy cows, with
633 special emphasis on coagulase-negative staphylococci. J Dairy Res **96:**1499-1510.
634 51. **Osés SM, Diez AM, Melero B, Luning PA, Jaime I, Rovira J.** 2013. Characterization by
635 culture-dependent and culture-independent methods of the bacterial population of
636 suckling-lamb packaged in different atmospheres. Food Microbiol **36:**216-222.
637 52. **Leng J, Cheng YM, Zhang CY, Zhu RJ, Yang SL, Gou X, Deng WD, Mao HM.** 2012. Molecular
638 diversity of bacteria in Yunnan yellow cattle (*Bos taurs*) from Nujiang region, China. Mol Biol
639 Rep **39:**1181-1192.
640 53. **Dungan RS.** 2012. Use of a culture-independent approach to characterize aerosolized
641 bacteria near an open-freestall dairy operation. Environ Int **41:**8-14.
642 54. **Lawrence PK, Shanthalingam S, Dassanayake RP, Subramaniam R, Herndon CN, Knowles**
643 **DP, Rurangirwa FR, Foreyt WJ, Wayman G, Marciel AM, Highlander SK, Srikumaran S.** 2010.
644 Transmission of *Mannheimia haemolytica* from domestic sheep (*Ovis aries*) to bighorn sheep
645 (*Ovis canadensis*): Unequivocal demonstration with green fluorescent protein-tagged
646 organisms. J Wildl Dis **46:**706-717.
647 55. **Blackall PJ, Bojesen AM, Christensen H, Bisgaard M.** 2007. Reclassification of [*Pasteurella*]
648 *trehalosi* as *Bibersteinia trehalosi* gen. nov., comb. nov. Int J Syst Evol Microbiol **57:**666-674.

649 56. **Norskov-Lauritsen N, Kilian M.** 2006. Reclassification of *Actinobacillus*
650 *actinomycetemcomitans*, *Haemophilus aphrophilus*, *Haemophilus paraphrophilus* and
651 *Haemophilus segnis* as *Aggregatibacter actinomycetemcomitans* gen. nov., comb. nov.,
652 *Aggregatibacter aphrophilus* comb. nov and *Aggregatibacter segnis* comb. nov., and
653 emended description of *Aggregatibacter aphrophilus* to include V factor-dependent and V
654 factor-independent isolates. Int J Syst Evol Microbiol **56:**2135-2146.
655 57. **Smith CB, Golden CA, Kanner RE, Renzetti AD.** 1976. *Haemophilus influenzae* and
656 *Haemophilus parainfluenzae* in chronic obstructive pulmonary disease. Lancet **1:**1253-1255.
657 58. **Kittelmann S, Kirk MR, Jonker A, McCulloch A, Janssen PH.** 2015. Buccal swabbing as a
658 noninvasive method to determine bacterial, archaeal, and eukaryotic microbial community
659 structures in the rumen. Appl Environ Microbiol **81:**7470-7483.
660 59. **Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, Davies J, Ervine A, Poulter L,**
661 **Pachter L, Moffatt MF, Cookson WOC.** 2010. Disordered microbial communities in asthmatic
662 airways. PLoS ONE **5:**e8578. DOI: 10.1371/journal.pone.0008578
663 60. **Bouljihad M, Leipold HW.** 1994. An ultrastructural study of pulmonary bronchiolar and
664 alveolar epithelium in sheep. J Vet Med A **41:**573-586.
665 61. **Griebel PJ, Entrican G, Rocchi M, Beskorwayne T, Davis WC.** 2007. Cross-reactivity of mAbs
666 to human CD antigens with sheep leukocytes. Vet Immunol Immunopathol **119:**115-122.
667 62. **Abraham W, Gill A, Ahmed A, Sielczak M, Lauredo I, Botinnikova Y, Lin K-C, Pepinsky B,**
668 **Leone D, Lobb R, Adams S.** 2000. A small-molecule, tight-binding Inhibitor of the integrin
669 α4β1 blocks antigen-induced airway responses and inflammation in experimental asthma in
670 sheep. Am J Respir Crit Care Med **162:**603-611.
671 63. **Abraham WM, Sielczak MW, Wanner A, Perruchoud AP, Blinder L, Stevenson JS, Ahmed A,**
672 **Yerger LD.** 1988. Cellular markers of inflammation in the airways of allergic sheep with and
673 without allergen-induced late responses. Am Rev Respir Dis **138:**1565-1571.

64. **Snibson KJ, Bischof RJ, Slocombe RF, Meeusen EN.** 2005. Airway remodelling and inflammation in sheep lungs after chronic airway challenge with house dust mite. Clin Exp Allergy **35:**146-152.

65. **Liravi B, Piedrafita D, Nguyen G, Bischof R.** 2015. Dynamics of IL-4 and IL-13 expression in the airways of sheep following allergen challenge. BMC Pulm Med **15:**101.

66. **Esther Jr CR, Boucher RC, Johnson MR, Ansede JH, Donn KH, O'Riordan TG, Ghio AJ, Hirsh AJ.** 2014. Airway drug pharmacokinetics via analysis of exhaled breath condensate. Pulm Pharmacol Ther **27:**76-82.

67. **Alton EWFW, Baker A, Baker E, Boyd AC, Cheng SH, Coles RL, Collie DDS, Davidson H, Davies JC, Gill DR, Gordon C, Griesenbach U, Higgins T, Hyde SC, Innes JA, McCormick D, McGovern M, McLachlan G, Porteous DJ, Pringle I, Scheule RK, Shaw DJ, Smith S, Sumner-Jones SG, Tennant P, Vrettou C.** 2013. The safety profile of a cationic lipid-mediated cystic fibrosis gene transfer agent following repeated monthly aerosol administration to sheep. Biomaterials **34:**10267-10277.

68. **El Mays TY, Choudhury P, Leigh R, Koumoundouros E, Van der Velden J, Shrestha G, Pieron CA, Dennis JH, Green FHY, Snibson KJ.** 2014. Nebulized perflubron and carbon dioxide rapidly dilate constricted airways in an ovine model of allergic asthma. Respir Res **15:**98.

69. **Popp JD, McAllister TA, Kastelic JP, Majak W, Ayroud M, VanderKop MA, Karren D, Yost GS, Cheng KJ.** 1998. Effect of melengestrol acetate on development of 3-methylindole-induced pulmonary edema and emphysema in sheep. Can J Vet Res-Rev Can Rech Vet **62:**268-274.

70. **Tsai L, Hoffman A, Berger R, Ingenito E.** 2004. Bronchoscopic lung volume reduction in a sheep model of heterogeneous emphysema. J Bronchology Interv Pulmonol **11:**83-86.

71. **Ingenito EP, Tsai L, Murthy S, Tyagi S, Mazan M, Hoffman A.** 2012. Autologous lung-derived mesenchymal stem cell transplantation in experimental emphysema. Cell Transplant **21:**175-189.

72. **Pohlmann JR, Akay B, Camboni D, Koch KL, Mervak BM, Cook KE.** 2012. A low mortality model of chronic pulmonary hypertension in sheep. J Surg Res **175:**44-48.

73. **McDonald JW, Ali M, Morgan E, Townsend ER, Cooper JD.** 1983. Thromboxane synthesis by sources other than platelets in association with complement-induced pulmonary leukostasis and pulmonary hypertension in sheep. Circ Res **52:**1-6.

74. **Weimann J, Ullrich R, Hromi J, Fujino Y, Clark MWH, Bloch KD, Zapol WM.** 2000. Sildenafil is a pulmonary vasodilator in awake lambs with acute pulmonary hypertension. Anesthesiology **92:**1702-1712.

75. **Julien M, Hoeffel JM, Flick MR.** 1986. Oleic-acid lung injury in sheep. J Appl Physiol **60:**433-440.

76. **Enkhbaatar P, Murakami K, Shimoda K, Mizutani A, Traber L, Phillips GB, Parkinson JF, Cox R, Hawkins H, Herndon D, Traber D.** 2003. The inducible nitric oxide synthase inhibitor BBS-2 prevents acute lung injury in sheep after burn and smoke inhalation injury. Am J Respir Crit Care Med **167:**1021-1026.

77. **Tsuno K, Prato P, Kolobow T.** 1990. Acute lung injury from mechanical ventilation at moderately high airway pressures. J Appl Physiol **69:**956-961.

78. **Gibbons MM, Dang X, Adkins M, Powell B, Chan P.** 2015. Finite element modeling of blast lung injury in sheep. J Biomech Eng **137:**041002. DOI: 10.1115/1.4029181

79. **Olivier A, Gallup J, De Macedo MMMA, Varga SM, Ackermann M.** 2009. Human respiratory syncytial virus A2 strain replicates and induces innate immune responses by respiratory epithelia of neonatal lambs. Int J Exp Pathol **90:**431-438.

80. **Morris DL, Clarkson MJ, Stallbaumer MF, Pritchard J, Jones RS, Chinnery JB.** 1985. Albendazole treatment of pulmonary hydatid cysts in naturally infected sheep: a study with relevance to the treatment of hydatid cysts in man. Thorax **40:**453-458.

81. **Enkhbaatar P, Joncam C, Traber L, Nakano Y, Wang J, Lange M, Connelly R, Kulp G, Saunders F, Huda R, Cox R, Schmalstieg F, Herndon D, Traber D.** 2008. Novel ovine model of

methicillin-resistant *Staphylococcus aureus*-induced pneumonia and sepsis. Shock **29:**642-649.

82. **Warner AE, Decamp MM, Molina RM, Brain JD.** 1988. Pulmonary removal of circulating endotoxin results in acute lung injury in sheep. Lab Invest **59:**219-230.

83. **Kramer BW, Kramer S, Ikegami M, Jobe AH.** 2002. Injury, inflammation, and remodeling in fetal sheep lung after intra-amniotic endotoxin. Am J Physiol Lung Cell Mol Physiol **283:**L452-L459.

84. **Engel M, Nowacki RE, Boden P, Reiss L, Uhlig S, Reynaert N, Gopal P, Wouters EM, Willems CMP, Kloosterboer N, Wolfs TAM, Zimmermann LI, Vos G, Kramer B.** 2015. The effects of dexamethasone and oxygen in ventilated adult sheep with early phase acute respiratory distress syndrome. Lung **193:**97-103.

85. **Langer T, Vecchi V, Belenkiy SM, Cannon JW, Chung KK, Cancio LC, Gattinoni L, Batchinsky AI.** 2014. Extracorporeal gas exchange and spontaneous breathing for the treatment of acute respiratory distress syndrome: An alternative to mechanical ventilation? Crit Care Med **42:**E211-E220. DOI: 10.1097/ccm.0000000000000121

86. **Bégin R, Rola-Pleszczynski M, Massé S, Lemaire I, Sirois P, Boctor M, Nadeau D, Drapeau G, Bureau MA.** 1983. Asbestos-induced lung injury in the sheep model: The initial alveolitis. Environment International **30:**195-210. DOI: 10.1016/0013-9351(83)90180-9

87. **Lee TC, Gold LI, Reibman J, Aston C, Bégin R, Rom WN, Jagirdar J.** 1997. Immunohistochemical localization of transforming growth factor-β and insulin-like growth factor-I in asbestosis in the sheep model. Int Arch Occup Environ Health **69:**157-164.

88. **Sebastien P, Begin R, Masse S.** 1990. Mass, number and size of lung fibers in the pathogenesis of asbestos in sheep. J Exp Pathol **71:**1-10.

89. **Palmarini M, Fan H, Sharp JM.** 1997. Sheep pulmonary adenomatosis: A unique model of retrovirus associated lung cancer. Trends Microbiol **5:**478-483.

90. **Leroux C, Girard N, Cottin V, Greenland T, Mornex JF, Archer F.** 2007. Jaagsiekte sheep retrovirus (JSRV): from virus to lung cancer in sheep. Vet Res **38:**211-228.

91. **Jorth P, Staudinger Benjamin J, Wu X, Hisert KB, Hayden H, Garudathri J, Harding Christopher L, Radey Matthew C, Rezayat A, Bautista G, Berrington William R, Goddard Amanda F, Zheng C, Angermeyer A, Brittnacher Mitchell J, Kitzman J, Shendure J, Fligner Corinne L, Mittler J, Aitken Moira L, Manoil C, Bruce James E, Yahr Timothy L, Singh Pradeep K.** 2015. Regional isolation drives bacterial diversification within cystic fibrosis lungs. Cell Host Microbe **18:**307-319.

92. **Ingenito EP, Solway J, McFadden ER, Pichurko B, Bowman HF, Michaels D, Drazen JM.** 1987. Indirect assessment of mucosal surface temperatures in the airways - theory and tests. J Appl Physiol **63:**2075-2083.

93. **Martin CJ, Cline F, Marshall H.** 1953. Lobar alveolar gas concentrations - effect of body position. J Clin Invest **32:**617-621.

94. **West JB.** 1978. Regional differences in the lung. Chest **74:**426-437.

95. **Robertson HT, Hlastala MP.** 2007. Microsphere maps of regional blood flow and regional ventilation. J Appl Physiol **102:**1265-1272.

96. **Glenny RW.** 2009. Determinants of regional ventilation and blood flow in the lung. Intensive Care Med **35:**1833-1842.

97. **Marsland BJ, Gollwitzer ES.** 2014. Host-microorganism interactions in lung diseases. Nat Rev Immunol **14:**827-835.

98. **Duncan SH, Louis P, Thomson JM, Flint HJ.** 2009. The role of pH in determining the species composition of the human colonic microbiota. Environ Microbiol **11:**2112-2122.

99. **Roth RR, James WD.** 1988. Microbial ecology of the skin. Annu Rev Microbiol **42:**441-464.

100. **Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW.** 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol **12:**87.

101. **Jones FS.** 1922. The source of the microorganisms in the lungs of normal animals. J Exp Med **36:**317-328.

102. **Ward DV, Gevers D, Giannoukos G, Earl AM, Methe BA, Sodergren E, Feldgarden M, Ciulla DM, Tabbaa D, Arze C, Appelbaum E, Aird L, Anderson S, Ayvaz T, Belter E, Bihan M, Bloom T, Crabtree J, Courtney L, Carmichael L, Dooling D, Erlich RL, Farmer C, Fulton L, Fulton R, Gao HY, Gill JA, Haas BJ, Hemphill L, Hall O, Hamilton SG, Hepburn TA, Lennon NJ, Joshi V, Kells C, Kovar CL, Kalra D, Li K, Lewis L, Leonard S, Muzny DM, Mardis E, Mihindukulasuriya K, Magrini V, O'Laughlin M, Pohl C, Qin X, Ross K, Ross MC, Rogers YHA, et al.** 2012. Evaluation of 16S rDNA-based community profiling for human microbiome research. PLoS ONE **7:**e39315. DOI: 10.1371/journal.pone.0039315

103. **Schloss PD, Gevers D, Westcott SL.** 2011. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS ONE **6:**e27310. DOI: 10.1371/journal.pone.0027310

104. **Nelson MC, Morrison HG, Benjamino J, Grim SL, Graf J.** 2014. Analysis, optimization and verification of Illumina-generated 16S rRNA gene amplicon surveys. PLoS ONE **9:**e94249. DOI: 10.1371/journal.pone.0094249

105. **Pinto AJ, Raskin L.** 2012. PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. PLoS ONE **7:**e43093. DOI: 10.1371/journal.pone.0043093

106. **Kembel SW, Wu M, Eisen JA, Green JL.** 2012. Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. PLoS Comput Biol **8:**e1002743. DOI: 10.1371/journal.pcbi.1002743

800 **Figure Legends**

801 **Fig 1: Diagram of the sheep lung, divided into anatomical segments:** Boxes indicate the
802 segments from which lung protected specimen brushings were taken in six sheep at three
803 time-points; these correspond to the RA: right apical, RCD: right caudal diaphragmatic, and
804 LCD: left caudal diaphragmatic segments.

805 **Fig 2: Location of brushings within the sheep lung:** Protected specimen brushings were
806 taken from the sections of the lung labelled A1-A9 and A13-A19 in one sheep at one time-
807 point.

808 **Fig 3**: **qPCR of lung brushings and controls**: The bronchoscope channel was flushed
809 through with 7.5 ml PBS and the wash collected (Wash Control n=18) prior to protected
810 specimen lung brushings being taken from sheep (Lung Brushings n=54). DNA was extracted
811 from Wash Controls and Lung Brushings and the quantity of bacterial DNA calculated using
812 16S rDNA qPCR. Lung Brushings were found to contain significantly higher quantities of
813 bacterial DNA than Wash Controls (Mann-Whitney U test: $P < 0.0001$). Negative controls
814 consisted of either water (n=3) or extraction kit controls (n=2). Boxes indicate interquartile
815 ranges and outliers are represented as diamonds.

816 **Fig 4: Clustering of time-points by lung microbiota composition:** PCOA graph showing
817 the similarities between bacterial communities sampled from three sheep lung segments in six
818 sheep at three time-points. Samples were found to cluster significantly by the time-point at
819 which they were taken (AMOVA: $P < 0.001$).

820 **Fig 5: The bacterial communities found in three separate lung segments within two**
821 **sheep**: Protected specimen brushings were taken from the lungs of sheep at three different
822 lung segments (RA: right apical, RCD: right caudal diaphragmatic and LCD: left caudal
823 diaphragmatic) at day 0. Sheep A (2S066) had highly different bacterial communities at each

824 lung segment whereas Sheep B (2D644) had similar bacterial communities at all three lung
825 sites.

826 **Fig 6: Clustering of individuals by lung microbiota composition**: PCOA graphs showing
827 the similarities between the bacterial communities extracted from protected specimen
828 brushing samples taken from sheep lungs at three time-points (baseline (0 days), one month
829 and three months). Samples were taken from three separate lung segments (RA: right apical,
830 RCD: right caudal diaphragmatic and LCD: left caudal diaphragmatic). Samples from within
831 the same sheep were found to cluster significantly at baseline (AMOVA: $P = 0.001$) and three
832 months (AMOVA: $P = 0.045$) but not at one month. This is likely to be due to the presence of
833 contaminants originating from the extraction kits in the one month samples.

834 **Fig 7: Diagram of the bacterial orders found in the sheep lung:** Bacterial orders found in
835 protected specimen brushings from the lung and trachea (A1-A9 and A13-A19), throat swab
836 and an extraction kit control taken during a study of one sheep at one time-point.

837 **Fig 8: Clustering of lung brushings by depth within the lung:** PCOA graph showing the
838 similarity of samples taken at different lung depths based upon the bacterial orders present.
839 Lung depths are represented by colour and correspond to different distances from the glottis
840 progressing in a caudal direction. Adjacent pairs of samples are represented by the same
841 symbol and colour. For the exact location of each sampling site see **Fig 2**. Sub-tracheal
842 samples ($\geq$ A3) clustered significantly by lung depth (AMOVA: $P= 0.033$) as did paired
843 samples (AMOVA: $P = 0.046$).

844

845 **Tables**

846 **Table 1: Sheep used in this study**

Sheep ID	Gender	Mean weight (kg) ± SD	Mean rectal temperature (°C) ± SD
2D618	Female	51 ± 3.1	39.0 ± 0.06
2S066	Male (castrated)	69 ± 2.6	39.6 ± 0.20
2D619	Female	59 ± 1.7	39.3± 0.20
2D620	Female	64 ± 4.6	39.1 ± 0.21
2D644	Female	65 ± 1.0	39.3 ± 0.06
2D645	Female	70 ± 2.0	39.4 ± 0.06

847

848 **Table 2**

849 Proportion of DNA sequence reads belonging to bacterial members of a mock community

Taxonomy	Expected proportion of reads	Actual proportion of reads (undiluted)	Actual proportion of reads (1:100)	Mock community species
Order				
Deinococcales	5%	24.48%	7.65%	*Deinococcus radiodurans*
Campylobacterales	5%	22.05%	12.65%	*Helicobacter pylori*
Bacteroidales	5%	19.59%	10.91%	*Bacteroides vulgatus*

Bacillales	20%	8.60%	22.40%	*Bacillus cereus, Listeria monocytogenes, Staphylococcus aureus, Staphylococcus epidermidis*
Lactobacillales	25%	5.10%	12.29%	*Enterococcus faecalis, Lactobacillus gasseri, Streptococcus agalactiae, Streptococcus mutans, Streptococcus pneumoniae*
Clostridiales	5%	4.01%	7.86%	*Clostridium beijerinckii*
Rhodobacterales	5%	3.92%	5.29%	*Rhodobacter sphaeroides*
Pseudomonadales	10%	3.42%	5.97%	*Acinetobacter baumannii, Pseudomonas aeruginosa*
Enterobacteriales	5%	3.33%	5.52%	*Escherichia coli,*
Neisseriales	5%	2.17%	3.49%	*Neisseria meningitidis*
Actinomycetales	10%	1.27%	2.92%	*Actinomyces odontolyticus, Propionibacterium acnes*
Other/unclassified	0%	2.03%	3.08%	
Genus*				
Deinococcus	5%	24.33%	7.61%	*Deinococcus radiodurans*
Helicobacter	5%	22.04%	12.65%	*Helicobacter pylori*
Bacteroides	5%	19.59%	10.90%	*Bacteroides vulgatus*
Rhodobacter	5%	3.91%	5.29%	*Rhodobacter sphaeroides*
Clostridium	5%	3.73%	7.59%	*Clostridium beijerinckii*
Staphylococcus	10%	3.04%	7.58%	*Staphylococcus aureus, Staphylococcus epidermidis*

Lactobacillus	5%	2.77%	6.59%	*Lactobacillus gasseri*
Pseudomonas	5%	2.33%	3.70%	*Pseudomonas aeruginosa*
Neisseria	5%	2.15%	3.27%	*Neisseria meningitidis*
Enterococcus	5%	1.40%	2.63%	*Enterococcus faecalis*
Acinetobacter	5%	0.97%	1.62%	*Acinetobacter baumannii*
Propionibacterium	5%	0.76%	1.77%	*Propionibacterium acnes*
Actinomyces	5%	0.48%	1.12%	*Actinomyces odontolyticus*
Streptococcus	15%	0.47%	1.63%	*Streptococcus agalactiae*, *Streptococcus mutans*, *Streptococcus pneumoniae*
Other/unclassified	0%	12.03%	26.05%	

* The species *Bacillus cereus*, *Escherichia coli* and *Listeria monocytogenes* could not be classified to genus level

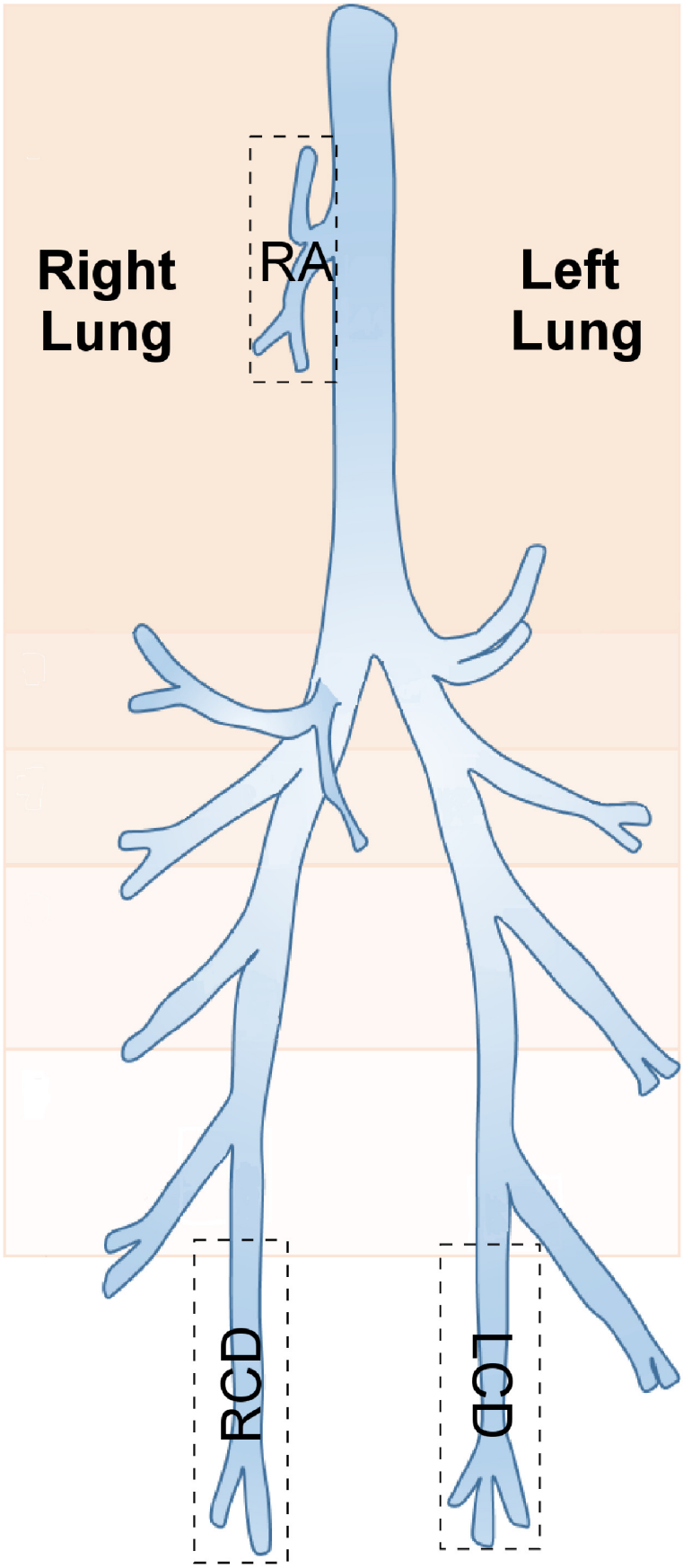
Table 3

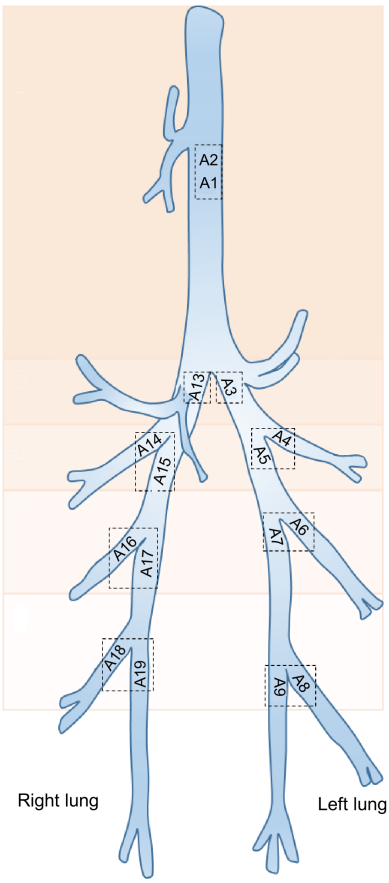
Abundance of the OTUs within the Pasteurellaceae family found in different locations of the sheep respiratory tract

Group	OTU 5: *Mannheimia* (%)	OTU 6: Pasteurellaceae (%)	OTU 7: *Bibersteinia* (%)	OTU 9: *Bibersteinia trehalosi* (%)
Throat swab	23.7	10.1	1.8	0.7
Trachea				
A1	5.5	4.5	28.4	5.3
A2	0	0.03	0	0.01

Left Lung				
A3	0	3.04	0.01	0.03
A4	1.2	0.006	0	0
A5	0	0.2	0.006	0
A6	0	0	0	0
A7	0.7	1.4	0.006	0.006
A8	0	0	0.8	0
A9	0.006	0.02	0.2	0
Right Lung				
A13	0.9	0	0	0.006
A14	0.006	0.3	0	0
A15	0	0.6	0.006	0.006
A16	2.3	0.6	0.006	0
A17	0.10	0	1.3	0.01
A18	0	0.02	0.01	0
A19	3.2	0	1.3	0

856





A1: Ventral trachea just caudal to bifurcation with right apical (RA) lobe segmental bronchus

A2: Dorsal trachea just caudal to bifurcation with right apical (RA) lobe segmental bronchus

A3: Medial aspect of left main bronchus immediately caudal to carina

A4: Medial aspect of first left ventral diaphragmatic segmental bronchus (LVD1), just caudal to bifurcation from left main bronchus

A5: Left main bronchus, lateral wall just after bifurcation to LVD1

A6: Medial aspect of second left ventral diaphragmatic segmental bronchus (LVD2), just caudal to bifurcation from left main bronchus

A7: Left main bronchus, lateral wall just after bifurcation to LVD2

A8: Third left ventral diaphragmatic segmental bronchus (LVD3)

A9: Left caudal diaphragmatic (LCD) segmental bronchus

A13: Medial aspect of right main bronchus immediately caudal to carina

A14: Medial aspect of first right ventral diaphragmatic segmental bronchus (RVD1), just caudal to bifurcation from right main bronchus

A15: Right main bronchus, lateral wall just after bifurcation to RVD1

A16: Medial aspect of second right ventral diaphragmatic segmental bronchus (RVD2), just caudal to bifurcation from right main bronchus

A17: Right main bronchus, lateral wall just after bifurcation to RVD2

A18: Third right ventral diaphragmatic segmental bronchus (RVD3)

A19: Right caudal diaphragmatic (RCD) segmental bronchus